

## ORIGINAL ARTICLE

## Pitch Range Variations Improve Cognitive Processing of Audio Messages

Emma Rodero<sup>1</sup>, Robert F. Potter<sup>2</sup>, & Pilar Prieto<sup>3,4</sup>

1 Department of Communication, Universitat Pompeu Fabra (UPF), 08108 Barcelona, Spain

2 The Media School, Institute for Communication Research, Indiana University, Bloomington, IN 47405, USA

3 Department of Translation and Language Sciences, Universitat Pompeu Fabra (UPF), 08108 Barcelona, Spain

4 ICREA. 08010, Barcelona, Spain

*This study explores the effect of different speaker intonation strategies in audio messages on attention, autonomic arousal, and memory. In this study, participants listened to 16 radio commercials, produced to vary in pitch range across sentences. Dependent variables were self-reported effectiveness and adequacy, psychophysiological arousal and attention, immediate word recall and recognition of information. Results showed that messages conveyed with pitch variations achieved better scores compared to commercials with a homogenous pitch range across the sentences. This was especially the case when high pitch intonation was followed by low pitch within a sentence. The results increase our understanding of the influence of pitch range on processing by establishing a concrete strategy as a best practice for improving attention and memory.*

**Keywords:** Pitch Range, Arousal, Attention, Immediate Recall, Recognition.

doi:10.1111/hcre.12109

People receive oral information everyday, through the electronic media or in interpersonal interactions. In all of these situations, the verbal message is important, but so is the nonverbal language, which supports the meaning of the message. *How* a speaker says something is as relevant as *what* he or she says (Tracy et al., 2011).

In terms of nonverbal communication, pitch can be considered as one of the most salient human communication features, because it conveys both semantic and pragmatic meaning (Perrachione, Fedorenko, Vinke, Gibson, & Dilley, 2013). It also represents a clear acoustic signal in the environment (Oxenham, 2012). Consequently, pitch is essential in auditory perception and necessary to understand and process messages (Cutler, Dahan, & Van Donselaar, 1997; Van Lancker, 1980). Previous work has focused on the influence of pitch on cognitive processing by manipulating pure or complex tones (see Lee, 2015, for a review). Scholars have done very little work on language, and that work focuses on isolated words or sentences, not on the entire

---

Corresponding author: Emma Rodero; e-mail: emma.rodero@upf.edu

message (Cutler et al., 1997). Also, a majority of studies have analyzed emotional intonation more than linguistic intonation (Tracy et al., 2011). Absent from the literature, however, are: (a) studies analyzing pitch range in complex communicative messages, presented in a sequence (see Lee, 2015); (b) research studying the concrete linguistic intonation patterns commonly used in mass media communications; and (c) studies about pitch range using psychophysiological measures to index cognitive processing. This study attempts to address these holes in the literature with an experiment designed to systematically analyze the impact of pitch range variations in radio commercials using a variety of dependent measures, including self-reported perceptions, immediate and delayed memory, and psychophysiological indicators of cognitive processing.

Regarding radio messages, we know that some structural features of the message, such as aspects of the announcer's voice, have an important impact on the cognitive processing of the listener, as they can provoke an automatic allocation of the limited resources in the human cognitive system (e.g., Lang, 2000; Potter, 2000). According to the limited capacity model of motivated mediated message processing (LC4MP; Lang, 2000, 2006), cognitive resources are allocated to messages in the form of both controlled and automatic processing (Lang, 2000). Controlled allocation of resources occurs as result of the information being of interest to the listener. For example, when a person hears a commercial for a product that they want to buy soon, they allocate controlled resources to determine whether the ad provides important information for their purchasing decisions. In addition to this, aspects of the audio production — such as voice changes and production effects — elicit automatic responses from listeners due to the auditory novelty introduced (Potter, 2000; Potter, Lynch, & Krause, 2015). The primary goal of this article is to study whether the pitch range of announcer speech can also enhance attention and memory in listeners.

To accomplish this, we employed different announcer intonation strategies in radio commercials to see how they affect the listener's cognitive processing. It is important to understand the effect of intonation strategy on audience cognition because, unlike other structural features such as jingles, sound effects, and multiple speakers, it is truly employable by any speaker creating an auditory message — recorded or live. In this study, we also employ a broad triangulation of measures to establish how pitch variations in a discourse affect self-reported message effectiveness and adequacy, autonomic arousal, cognitive resource allocation, immediate word recall, and sentence recognition.

### **Intonation as a dual mode of meaning**

Intonation is conceptually defined in this study as a spoken sequence of pitch variations (high tones and low tones) that functions to group syntactic units, as well as reflecting pragmatic meanings related to information structure and a wide variety of linguistic phenomena (e.g., Globerson, Amir, Golan, Kishon-Rabin, & Lavidor, 2013; Hammerschmidt & Jurgens, 2007; Wells, 2007). As Wells (2007) explains, intonation is the melody of speech. This intonation melody is produced through

the combination of two general levels, high (H) and low (L) pitch, according to the autosegmental–metrical (AM) model (Ladd, 1996; Pierrehumbert, 1980). In English, these two tones are used to distinguish boundary tones and pitch accents. The rest of the variations, including pitch range changes, usually convey different levels of emphasis (Dilley, 2010), and in certain cases, categorical differences in meaning (Borràs-Comes, del Mar Vanrell, & Prieto, 2014; Braun, 2006). In this study, we do not use categorical meaning. The most common combination in a declarative sentence is high pitch followed by low pitch during the duration of the sentence (Hirschberg, 2003; Ladd, 1996). In this pattern, high pitch may operate to attract the listener's attention and then low pitch and descending contours may convey the full extent of the message. Thus, the way pitch range variations occur across the sentence might be particularly influential during the processing of the information, as a dual mode of meaning.

Some authors have demonstrated that intonation influences the cognitive processing of messages (e.g., Hirschberg & Pierrehumbert, 1986; Levi & Pisoni, 2007; Rodero, 2015), but few studies have focused on pitch range. Fraundorf, Watson, and Benjamin (2010) analyzed the effects of pitch accenting and showed that memory tests using probes accented with a high pitch range increased hits to correct statements and correct rejections of foils without impairing memory for other parts of the discourse. These results suggest that listeners use pitch range differences to encode and update discourse information and thus facilitate memory and recall.

Recent research suggests that the intonation strategies of broadcasters are not optimal for perception and cognitive processing in audiences (Bolinger, 1998; Francuz, 2010; Nihalani & Po Lin, 1998; Price, 2008; Rodero, 2013, 2015; Strangert, 1991). However, none of these studies have investigated the effects of pitch range variations across the sentence as an intonation strategy in an entire message. We used text from actual radio advertisements to produce our stimuli varying in pitch range to test our hypotheses. Furthermore, in addition to self-report and memory data, we use psychophysiological indices of cognitive processing and autonomic arousal.

### **Pitch variation and the dual function**

Gussenhoven (2004) establishes two types of pitch variations in the curve of intonation: pitch level and pitch range. Pitch level represents the register or midline where the melodic contour is delivered while pitch range is the difference or distance between the highest and the lowest pitch level. In this study, pitch variations were produced by varying both pitch level and range. We created radio commercials to include one of four intonation patterns for comparison. In two commercials, the announcers spoke homogeneously in either the expanded pitch range (HH, High) or the compressed pitch range (LL, Low). The HH pattern was produced with a high fundamental frequency (F0) topline and homogeneous expanded pitch range of most pitch accents, and the LL pattern was produced with a low F0 topline together with homogeneous compressed pitch range in most pitch accents. These are the most common patterns used in speaking in public and on the radio and television,

both news and commercials (Bolinger, 1998; Cotter, 1993; De-la-Mota & Rodero, 2011; Francuz, 2010; Grawunder, Engert, Kaiser, Unger, & Bose, 2008; McGregor & Palethorpe, 2008; Price, 2008; Rodero, 2015). The HH is the most typical in media messages, and it is suggested as an announcing practice (Neil, Worrall, Day, & Hickson, 2003; Stephenson, Reese, & Beadle, 2013). The HH is a very emphatic pattern used by broadcasters for attracting attention. The LL pattern, more monotonous, can be found in public speaking (not trained speakers) and in some commercials for conveying relax or sensuality (Prieto, Payà, & Vanrell, 2014). In the other two levels, the announcer combined the two pitch ranges across two pitch variations: the most typical in declaratives: High-Low; and the opposite, Low-High (henceforth HL and LH, respectively).

A central hypothesis of this study is that exploiting pitch range variations in different parts of the sentences will affect the listener's cognitive processing of the message. This hypothesis is based on the principle of distinctive and contrastive coherence of prosody (Rodero, 2015). While the distinctive function distinguishes different parts of the sentence into prosodically distinct units, the contrastive function establishes an acoustic contrast between high and low pitch. Consequently, pitch variations fulfill a dual function. Rodero (2015) applied this principle to two models of radio news bulletins with different intonation/stress and speech rates. The results showed that the recognition of information was highest for news bulletins where this principle was observed. For this reason, we expect that stimuli in the current study featuring pitch change will result in better processing than stimuli with homogenous pitch. However, because both the HL and the LH patterns follow the principle of distinctive and contrastive coherence, it is important to consider which combination is likely to be optimal.

A pitch range contrast prosodically may emphasize both syntactic and semantic distinctions between the two parts of the sentence. The syntax plays a key role in this study, as it is important to choose a concrete pitch pattern (Selkirk, 2008). The syntactic distinction is the part before the variation being recognized as distinct from the part after it. However, the semantic differentiation is one of relative meaning among the topics in the discourse. In English, the most common semantic divisions put "accessory information" in the first part of a sentence and more relevant information in the second. For example, in the statement "*If you have problems in Mathematics*, Potter's Learning Center can help," having a pitch change occur after the italicized segment sets the two phrases apart both syntactically and semantically. Once again, syntactically, the change in pitch communicates that two separate ideas are present. Most commonly, English listeners receive accessory information at the beginning of the sentence with more important information following (Bock & Mazzella, 1983). Furthermore, it is more prevalent in English that accessory information is delivered in a high pitch and then elaborated upon with the most important details communicated in a low pitch (Estebas-Vilaplana, 2008; Gussenhoven, 2004; Ladd, 1996). The higher the initial pitch, the stronger the break with the preceding sentence will be

perceived (Gussenhoven, 2004). In other words, although some pitch variation is better than none due to the automatic attention capture that happens at the point of the pitch change (Potter, Jamison-Koenig, Lynch, & Sites, 2016), the more recalled pattern in English would be HL. Also, as some studies have shown, prosody as a form of intonation provides the basic auditory structure to maintain an acoustic sequence in memory while the brain processes it (Frazier, Carlson, & Clifton, 2006). The HL pattern should be easier retained in memory as contrastive elements are better recalled.

The LH pattern reverses this optimal order, the important information in the last part of the sentence is delivered in a high pitch while “accessory” information at the beginning is identified with a low pitch. Potter et al. (2016) have shown that unexpected changes in pitch result in worse memory than expected changes. Therefore, the following hypothesis is made:

H1: The patterns implementing the HL pitch range variation will attain the highest level of recognition memory and immediate recall compared to the other patterns.

In the current study, furthermore, we predict that psychophysiological correlates of effective cognitive processing are likely to be greater for messages containing pitch variations than those containing homogenous pitch. However, similar to our rationale for the first hypothesis, we feel that there will be greater arousal and lower heart rate (HR; indicative of increases cognitive resource allocation) in response to the HL messages compared to all the others. High-pitched sounds, due to their acoustic characteristics, attract attention, which is why we produce alarm claims in a high pitch. We can observe the same principle in voice perception. The high voice superiority effect (e.g., Marie & Trainor, 2013) shows the perceptual dominance of sounds with high F0; a phenomenon that research has observed in humans as early as 3 months of age. The physiological properties of the cochlea, which are designed to respond quickly to high pitch sounds (Marie & Trainor, 2013), might explain this perceptual effect. Moreover, expanded pitch range has been correlated with emphasis (Lieberman & Pierrehumbert, 1984). Consequently, high pitch can function to grab attention (Hill & Miller, 2010). Then, once captured, the most important part of the sentence is delivered in a low pitch to conclude the message (Francuz, 2010; Pierrehumbert & Hirschberg, 1990; Rodero, 2015). This contrastive effect should attract attention and increase activation. Because the LH pattern also involves a contrast between two level tones (Low-High), we might hypothesize that this pattern could also increase attention and arousal levels. However, because there is a strong tendency for declarative intonation patterns to be HL, we also expect that LH pattern will achieve less arousal and cognitive resource allocation than the HL pattern. The second hypothesis follows:

H2: The HL pattern of pitch range variation will elicit the greatest levels of sympathetic nervous system activation (measured as skin conductance) and cognitive resource allocation (measured as HR) compared to the other patterns.

Finally, we make predictions about how the stimuli will be perceived by subjects regarding their effectiveness and adequacy of audio messages. We predict that the

messages with pitch variations will be better perceived than those without because people are used to listening to different pitch variations. In fact, when someone talks, they use different tones and not a constant or repetitive intonation. Rodero (2013) showed that the repetitive intonation pattern used by broadcasters was negatively assessed compared to a natural intonation pattern. As pitch variations sound more natural, the participants should rate this pattern as more adequate. However, the HL pattern is hypothesized to be the most effective and adequate. From a linguistic point of view, the HL pattern exploits a persuasive characteristic of the intonation of broad focus statements across languages (e.g., Cruttenden, 1997; Grosz & Sidner, 1986; Hirschberg & Pierrehumbert, 1986; Ladd, 1996; Wells, 2007). This is a common intonation used in declaratives in many languages. Therefore, the final hypothesis is posited:

H3: The HL pattern will result in the highest self-reported effectiveness and adequacy compared to the rest of the patterns.

## Method

The design of the experiment was a 4 (intonation pattern)  $\times$  4 (message)  $\times$  4 (order) mixed factorial design. The order was a between-subject factor. The intonation pattern was a within-subject factor, with four levels. The first intonation pattern was delivered in a consistently low pitch range, the second pattern in a consistently high pitch range, the third began with high pitch and transitioned to low pitch, and the fourth began with low pitch and transitioned to high pitch. Each pattern of intonation consisted of two informative and two narrative messages. Each intonation pattern was in all the possible positions. Participants were randomly assigned to an order.

## Participants

Fifty-one participants (30 females) from a university in the Midwestern United States participated in the main study. Most of the participants (90%) were native to North America, but all had a high level of English proficiency. The average age was 31 years.

## Procedure

After obtaining informed consent, physiological sensors were attached to the palm and forearms of the participant. Each participant then listened to the 16 commercials using quality headphones. Immediately after each message, the participant was given 30 seconds to complete the immediate recall task. They then had to assess the commercial using the effectiveness and adequacy scales. This same procedure was repeated with all 16 commercials. The sensors then were removed and, as a distraction task, they had to complete a test of resting trait motivation activation by rating 41 still images on three 9-point scales (Lang, Kurita, Rubenking, & Potter, 2011). These data are not reported here. Finally, the participants completed the recognition test for the radio advertisements. Each participant received a \$30 gift card for their participation.

## Stimuli

Sixteen radio ads were selected from a pool of Mercury Award winning advertisements from the years 2000 to 2003. These were selected due to their being written, so that simple speech was the main element, expressed by only one voice, and with simple syntactic structures. All the ads selected offered products or services of interest to young people to ensure participant involvement and motivation. To control for two of the primary creative strategies, half the ads employed a narrative style and half an informative style. Transcriptions of the commercials were made to modify the texts by reducing the duration of each commercial from 1 minute to 25 seconds. Therefore, the final stimuli lasted from 21 to 25 seconds.

The edited scripts were then produced by a professional male announcer combining the best values for pitch and voice quality (Jones, Feinberg, DeBruine, Little, & Vukovic, 2008; Rodero, 2013). The announcer had a relatively low pitch ( $M = 120$  Hz,  $SD = 30$ ), with a resonant voice quality. We selected a male voice to test pitch variations because they are more common in radio commercials. The speaker was coached in how to produce the precise prosodic variations required in the experiment. For the recordings of the HL and the LH patterns, two different parts in each sentence had to be prosodically distinguished, namely subject and predicate (e.g., *Larson mowers/offer more bang for the buck than any other commercial mower*) and subordinate clause and main clause (e.g., *If there's one thing I remember about my father/it's the time we spent ice fishing*). Then the 16 ads were recorded in optimal acoustic conditions. The final version recorded by the announcer was not edited or manipulated.

The acoustics of the stimuli recordings were pretested using *Praat* software (Boersma & Weenink, 2016) to ensure they provided the patterns described in the experimental design. Results showed statistically significant differences for each prosodic pattern. First, there were significant differences in mean pitch level,  $F(3,93) = 6.124$ ,  $p < .001$ , among the four patterns of intonation: LL ( $M = 93$  Hz), the HH (144 Hz), the HL (115 Hz), and the LH patterns (116 Hz). In patterns with sequential variations of pitch range, there were significant differences between the HH (130 Hz) and the LL (101 Hz) portions of the sentence,  $F(3,93) = 241$ ,  $p < .001$ .

Participants then perceptually pretested the 16 stimuli. Participants were ( $N = 20$ ; 10 female) similar in age to those who would take part in the experiment. The pretest procedure consisted of an online survey where individual respondents listened to the ads and rated them using the effectiveness and adequacy scales. The goal of the pretest was to see if participants perceived the commercials as similar to those commonly heard on radio, and to confirm that they did not sound exaggerated or unnatural. The results confirmed that there were significant differences among the four patterns for effectiveness,  $F(3,17) = 8.80$ ,  $p = .004$ , and adequacy,  $F(3,17) = 10.34$ ,  $p < .001$ . The effectiveness index ( $M = 3.36$ ,  $SD = 1.23$ ) and the adequacy index ( $M = 3.65$ ,  $SD = 1.35$ ) were above the scale midpoint (2.50). Pretest results also showed that the different pitch variations were perceived according to the intended pattern.

### Dependent variables

Tests of immediate recall and recognition accuracy measured memory. First, immediate word recall was measured after listening to each ad. Participants had 30 seconds to type as many words from the ad as they recalled into a blank box appearing on the computer screen. After all the stimuli were heard, and the subject completed a simple distraction task, we measured recognition memory using a forced choice, yes/no, timed recognition test to assess the efficiency of encoding. In the recognition task, each subject listened to 48 isolated sentences, three for each commercial message. Each of these sentences was taken from a different part of the ad (introduction, middle, and conclusion). The participants had to recognize whether they had previously heard them in the original commercial. There were 24 targets and 24 foils created by replacing two of the actual words spoken with synonyms. Using the keyboard, subjects pressed one button if they had heard that segment earlier (Yes) and another button if they had not (No). We scored recognition accuracy as 1 for targets correctly identified.

We collected psychophysiological measures of arousal and attention, time-locked to the stimuli presentation. We operationalized arousal as the response of the sympathetic nervous system used to measure motivational response as well as an increase in engagement. Electrodermal activity (EDA), to measure arousal, was quantified by measuring the conductance of a constant voltage (.5 V) delivered to two 8 mm Ag–AgCl electrodes placed on the palmar surface of the nondominant hand. EDA values were generated by a Biopac (Biopac Systems, Inc., Goleta, CA) PPGEDR Bionomadix (Biopac Systems, Inc., Goleta, CA) skin conductance coupler sampled at 1,000 Hz and aggregated at the 1-second level. Statistical analyses were performed in the first 21 seconds of radio ads collapsed across ad and condition.

Cognitive resource allocation was operationalized using cardiac activity measured initially as the electrocardiogram (ECG). The ECG was measured during the stimuli presentation by means of two Beckman standard 8 mm Ag–AgCl electrodes (Beckman Instruments Inc., Fullerton, CA, USA) placed on the forearms with a ground electrode on the nondominant forearm. The ECG was sampled at 1,000 Hz using a Biopac bioamplifier. We converted milliseconds between the QRS-complex in the ECG into HR (in beats per minute [BPM]/second) and used as a physiological measure of attention, with lower BPM rates indicative of greater attention being paid to the radio messages (Potter & Bolls, 2012).

Self-reported effectiveness and adequacy were measured using 5-point scales to assess listener perception of the commercials. Listeners were asked how clear, natural, pleasant, dynamic, and persuasive the voice was perceived to be. Calculating the mean of all these variables created the effectiveness index. The adequacy index was comprised of the mean degree of correctness, appropriateness, comprehensibility, and credibility of the voice. These scales have been demonstrated as a valid in past research (Rodero, Larrea, & Vázquez, 2013) and achieved strong reliability (Cronbach's alpha coefficient for effectiveness = .892; for adequacy = .725) in the current study.<sup>1</sup>



## Results

We submitted the data to a 4 (patterns of intonation) by 4 (orders) general linear model (GLM) with a Latin-Square analysis. The physiological data were submitted to a 4 (intonation pattern) by 4 (message) by 21 (time) repeated measures analysis of variance (ANOVA) for statistical analysis. The time factor was determined by the duration of the shortest commercial (21 seconds).

### Effects of pitch variation on memory

The first hypothesis posited that the HL messages would result in greater memory than the other messages. We first tested this with the immediate recall memory, which we measured as the percentage of correctly recalled words that had been heard in the ads. There was a significant main effect for intonation pattern,  $F(3,816) = 8.39, p < .001$ . The pattern with the highest percentage of recalled words was HL ( $M = 21.48, SD = .63$ ), followed by LL ( $M = 20.68, SD = .59$ ), HH ( $M = 19.98, SD = .53$ ), and LH ( $M = 17.64, SD = .51$ ). Post hoc tests showed significant differences between the HL pattern and the rest. The recall results confirm this hypothesis.

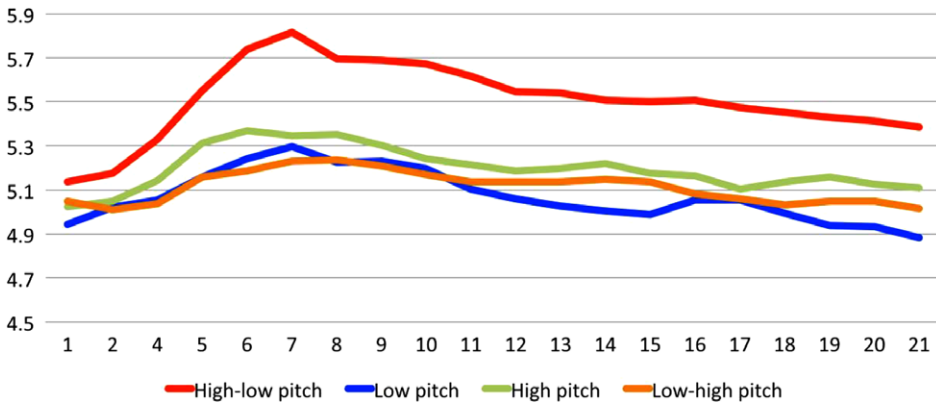
Memory was also tested with subsequent forced-choice recognition memory for targets and foils. There was a significant main effect for intonation pattern,  $F(3,48) = 12.69, p < .001$  for pattern. The patterns with the greatest level of recognition were those with variations: HL ( $M = .81, SD = .39$ ) followed by LH ( $M = .78, SD = .41$ ). The patterns with no variations had the lowest levels of recognition: HH ( $M = .69, SD = .46$ ) and LL ( $M = .68, SD = .46$ ). Post hoc results showed that the no-variations patterns, HH and LL, were statistically different from the variations patterns, HL and LH. Therefore, the first hypothesis of this study is only partially confirmed with the forced choice recognition data; although the HL recognition had the highest mean recognition data, it was only statistically significant from the homogenous pitch conditions.

### Effects of pitch variation on physiological correlates of arousal and attention

Hypothesis 2 proposed that the pattern of pitch range variation should elicit the greatest levels of sympathetic nervous system activation and cognitive resource allocation compared to the other patterns.

We found main effects for intonation pattern on average skin conductance level,  $F(3,48) = 2.70, p = .047$ , and a significant Pattern  $\times$  Time interaction,  $F(3,48) = 1.52, p = .006$ . The HL pattern elicited the highest skin conductance ( $M = 5.54, SD = .80$ ), followed by HH ( $M = 5.19, SD = .87$ ), LH ( $M = 5.11, SD = .79$ ), and LL ( $M = 5.10, SD = .77$ ). Posttest results showed significant differences between HL pattern and the rest. Figure 1 shows these data.

The results of HR indicated only significant main effects for Time,  $F(2,48) = 20.14, p < .001$ . Nevertheless, when collapsed across time, the mean HR values for each pitch pattern were in the predicted direction with the slowest HR—indicative of the greatest allocation of processing resources—measured during the HL ads. The HL ads



**Figure 1** Skin conductance levels.

( $M = 93.35$ ,  $SD = 2.99$ ) obtained the lowest levels followed by LH ads ( $M = 93.71$ ,  $SD = 2.89$ ), LL ( $M = 94.27$ ,  $SD = 2.92$ ) and HH ( $M = 95.05$ ,  $SD = 3$ ). Therefore, the second hypothesis of this study was confirmed for autonomic arousal, where HL was significantly greater than the rest. The means of the cardiac data were in the predicted direction — with lowest average HR for HL pitch variation compared to the rest — but failed to reach statistical significance.

### Effects of pitch variation on perceptions of effectiveness and adequacy

We also predicted that the HL pattern should achieve the best level of self-reported effectiveness and adequacy. We performed a general linear model (GLM) with Latin-Square analysis. Using an alpha level of .001 to evaluate homogeneity assumptions, Box's  $M$  test of homogeneity of covariance was significant ( $p = .028$ ) as well as Levene's homogeneity of variance for effectiveness ( $p < .001$ ) and for adequacy ( $p < .001$ ). As the principle of homoscedasticity was violated, we applied two tests to show whether or not differences between the groups were statistically significant. Both the Welch test and the Kruskal–Wallis  $H$  test were significant for effectiveness ( $p < .001$ ) and for adequacy ( $p < .001$ ). Table 1 shows the means and standard deviation for the effectiveness and adequacy measures. The order of presentation was not significant in any of the analyzed variables.

Regarding the effectiveness index, main effects were found for intonation pattern,  $F(3,811) = 7.25$ ,  $p < .001$ . The results indicated that the pattern HL obtained the best score, followed by LH, HH, and LL. Post hoc tests showed that the pattern HL was statistically different from the other patterns. There were no differences among the other three patterns.

Regarding the adequacy index, there was a main effect for pattern  $F(3,811) = 7.36$ ,  $p < .001$ . The HL pattern scored the highest, followed by LH, LL, and HH (Table 1). Posttest results showed that the no-variations patterns, HH and LL, were statistically different from the variations patterns, HL and LH.

**Table 1** Means and SD for Effectiveness and Adequacy

	Effectiveness		Adequacy	
	Mean	SD	Mean	SD
High-Low pitch	3.35	.92	3.48	.99
Low-High pitch	3.11	.86	3.45	.93
High pitch	3.04	.89	3.08	.98
Low pitch	2.95	.93	3.29	.91

Note: SD = standard deviation.

The results confirm the hypothesis for the effectiveness variable and partially for the adequacy variable, where HL was statistically different from the rest of the patterns.

## Discussion

The purpose of this study was to determine whether pitch range variations in audio messages, such as radio advertisements, would influence listener's perception of effectiveness and adequacy, physiological arousal and attention, and memory. As predicted by our theoretical model (LC4MP; Lang, 2000, 2006), different pitch strategies modified the cognitive processing of the ads on the part of the participants to a greater extent than more homogenous vocal strategies. The pattern High-Low pitch (HL) attained the most favorable results in perceptions of adequacy and effectiveness, autonomic arousal, and both immediate and delayed memory. The mean cardiac data was also suggestive of the superiority of having announcers vary the pitch of their delivery by starting high and changing to low. However, the results did not reach statistical significance.

These results might be explained by both linguistic and psychological reasons, in line with the principle of distinctive and contrastive coherence of prosody (Rodero, 2015).

First, about recognition and recall, the results showed that the best pattern was HL, particularly in recall. In line with the distinctive function of prosody (Rodero, 2015), the distribution of information between accessory and important data, signaled by an acoustic contrast, could have simplified the encoding process, so the listener still had available resources for storage of information, and thus recognition and immediate recall improved, according to the results of arousal. As a consequence, the cognitive processing of the information was attained, in line with the LC4MP (Lang, 2000, 2006). In the second pattern with pitch variations (LH), the level of recognition obtained the second best result, with no significant differences with HL. This finding also indicated that here pitch variations helped to distribute and to differentiate information, allocating more resources to encoding the message. Therefore, the encoding process was improved and the level of recognition increased. However, the immediate recall was negatively affected. In this variable, LH did

not obtain significant differences with no-variations patterns. Thus, for immediate recall, only the HL strategy correctly worked. This seems to suggest that the order of pitch variations was important for recall, as for arousal. The storage process needed not only pitch variations but also a high pitch at the beginning of the sentence for accessory information and to prepare listeners for the important data. The patterns with no variations achieved poor results due to the lack of acoustic differentiation between parts of the sentences and, therefore, because the principle of distinctive coherence of prosody was not observed (Rodero, 2015).

Secondly, regarding arousal and attention, the HL pattern also obtained the best results. Acoustic changes provoked by pitch variations contributed to attracting the listener's attention, as was shown by the lowest HR level (although with not significant differences) and to increase their arousal, as shown in EDA results. In line with the contrastive function of prosody (Rodero, 2015), the initial high pitch could have acted as an orienting element of attention (increasing the levels of attention and arousal) and then preparing the listener to hear the most important part. Although with no significant differences, the slowest HR—indicative of the greatest allocation of processing resources—was measured in HL. Furthermore, the increase in the activation allocated more resources to process the message and could indicate a motivational response. Therefore, this pattern achieved the best results and was different from the rest. The second pattern with variations (LH) did not obtain significant differences with no-variations patterns. This appears to indicate that the order of pitch variations was very important for arousal, as well as for recall. It was difficult to provoke a high level of arousal and/or attention when the sentences began with a low pitch, which was not grabbing the attention and renewing it in every utterance. Regarding the patterns with no-variations, the processing of information by the listener suffered because there was not acoustic differentiation between different parts of the sentences and, therefore, the principle of contrastive coherence of prosody was not observed (Rodero, 2015).

Third, with regard to self-reported perceptions of the stimuli, participants considered the HL pattern as the most effective and adequate. The reason is that this pattern used the most common model of intonation in declaratives sentences. Therefore, the participants could have perceived it as the most effective due to the familiarity effect. In addition, the acoustic contrast between high and low pitch might have sounded more natural and, thus, been perceived as the most adequate. The second pattern with variations did not achieve the best results due to its unusual structure (Low-High pitch), more proper for an interrogative than a declarative sentence. The patterns with no-variations achieved the worst results in self-perception, with no significant differences between them.

This study extends the current research about pitch range processing by applying the variations to a complex communicative message (e.g., radio ads), by studying the cognitive processing of the most common pitch strategies used in public speaking and in media messages, and by showing the effect of pitch range on the listener's physiological level. Also, if we consider that a key goal of any public speaking

communication is to maintain listener's attention and that these listeners understand the content of messages, it is clear that the results of this study may have significant repercussions for all the professionals related to speech. The conclusions may be of interest for the study of public speaking and its relationship with persuasive communication. The findings may be applied to improve the processing of auditory messages in media communication (audio in different platforms: radio, podcast, television and the Internet) and specifically in advertising and marketing. The application of persuasive strategies to improve the effectiveness of advertisements is a key aspect to increase the sale of products or services. Other potential areas could benefit from the results: political communication, education activities, language learning, or speech synthesis. Last, the conclusions of this study can be applied to health care to improve the cognitive processing of individuals with attention or recall disorders, for example, elder or children with attention deficit.

### **Limitations and future research**

Concerning the patterns employed, it is worthwhile considering that, in normal conversation or reading, people vary the intonation more than the simple pitch variations presented in this research and they do not use only one pitch pattern. But in this study intonation has been simplified to four patterns to control the analyzed variables. Hence new patterns, combining different pitch patterns, will be considered in future research.

Another important limitation was that participants listened to 16 radio commercials in only one session. This amount of commercials could have caused fatigue on the part of the listeners. Consequently, future research would try to reduce this amount of commercials to test whether or not the cognitive processing improves. In addition, the different structure of the ads, informative and narrative, did not obtain significant results. This could be due to the content of the selected commercials. Thus, this variable should be better controlled in future research.

Finally, this study employed only a male voice to test the different patterns. Pitch is a voice element very related to gender and that is why the same intonation patterns conveyed with a female voice might obtain different results. Therefore, it would also be appropriate to extend the study to female voices, which are usually more melodic.

### **Acknowledgments**

This research was supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme (FP7-PEOPLE-2012-IOF, 328636). Thanks to the Institute for Communication Research in The Media School at Indiana University where this research was conducted.

### **Appendix**

#### **Example of informative ad**

If you always wanted to get your MBA, we can help you. At Howard's MBA Institute, we've just slashed the entrance fee by 50%! Call before midnight to secure your space

in our MBA program. It just takes one phone call, one year, and one heck of a lot of determination! Mention this commercial and receive an additional 20% off at the campus bookstore. The number to call is 956-789-2457. See you this September!

### Example of narrative ad

Today, Meredith Chase will be preparing a simple Italian-style dinner. She begins by dicing fresh tomatoes and the top-left part of her knuckle. Bleeding profusely, she then begins to chop two large onions, causing tears to stream from her eyes. Behind her, the unattended pan burns everything to a crisp. Fortunately, Meredith had downloaded the Bay fire alarm app to alert the authorities. Don't be a victim. Bay fire alarm app in the Apple Store.

### Note

- 1 Furthermore, a factor analysis was conducted for the scale items. All the items of the scales were assessed on the appropriate factor (.5 or higher). Prior to performing principal component analysis (PCA), the suitability of the data for factor analysis was assessed. The Kaiser–Meyer–Olkin value was 0.71, exceeding the recommended value of 0.6, and Bartlett's Test of Sphericity attained statistical significance, supporting the factorability of the correlation matrix. PCA with Varimax rotation explained the greatest variance, at 76 and 61% for the two factors, respectively.

### References

- Bock, J. K., & Mazzella, J. R. (1983). Intonational marking of given and new information: Some consequences for comprehension. *Memory and Cognition*, **11**, 64–76. <https://doi.org/10.3758/BF03197663>
- Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer (version 5.3.04)*. Retrieved from <http://www.fon.hum.uva.nl/praat>
- Bolinger, D. (1998). The network tone of voice. *Journal of Broadcasting*, **26**, 726–728.
- Borràs-Comes, J., del Mar Vanrell, M., & Prieto, P. (2014). The role of pitch range in establishing intonational contrasts. *Journal of the International Phonetic Association*, **44**(1), 1. <https://doi.org/10.1017/S0025100313000303>
- Braun, B. (2006). Phonetics and phonology of thematic contrast in German. *Language and Speech*, **49**(4), 451–493. <https://doi.org/10.1177/00238309060490040201>
- Cotter, C. (1993). Prosodic aspects of broadcast news register. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, **19**, 90–100. <https://doi.org/10.3765/bls.v19i1.1520>
- Cruttenden, A. (1997). *Intonation*. New York, NY: Cambridge University Press.
- Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, **40**(2), 141–201.
- De-la-Mota, C., & Rodero, E. (2011). La entonación en la información radiofónica [Intonation in radio news]. *Quaderns de Filologia: Estudis de Comunicació*, **45**, 159–176.

- Dilley, L. C. (2010). Pitch range variation in English tonal contrasts: Continuous or categorical? *Phonetica*, *67*, 63–81. <https://doi.org/10.1159/000319379>
- Estebas-Vilaplana, E. (2008). Modelling final declarative intonation in English and Spanish. In M. T. Gibert & L. Alba (Eds.), *Estudios de filología inglesa: homenaje a la Dra. Asunción Alba Pelayo* (pp. 109–129). Madrid, Spain: UNED.
- Francuz, P. (2010). The impact of audio information intonation on understanding television news content. *Psychology of Language and Communication*, *14*(1), 71–86. <https://doi.org/10.2478/v10057-010-0005-6>
- Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how *contrastive* contrastive accenting really is. *Journal of Memory and Language*, *63*, 367–386. <https://doi.org/10.1016/j.jml.2010.06.004>
- Frazier, L., Carlson, K., & Clifton, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, *10*(6), 244–249. <https://doi.org/10.1016/j.tics.2006.04.002>
- Globerson, E., Amir, N., Golan, O., Kishon-Rabin, L., & Lavidor, M. (2013). Psychoacoustic abilities as predictors of vocal emotion recognition. *Attention, Perception and Psychophysics*, *75*(8), 1799–1810. <https://doi.org/10.3758/s13414-013-0518-x>
- Grawunder, S., Engert, H., Kaiser, H., Unger, S., & Bose, I. (2008). Pitch, speech rate and rhythm as speaking style characterizing parameters in German news presentation. In *EASR08*. London, England: UCL.
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, *12*(3), 175–204.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge, England: Cambridge University Press.
- Hammerschmidt, K., & Jurgens, U. (2007). Acoustical correlates of affective prosody. *Journal of Voice*, *21*(5), 531–540. <https://doi.org/10.1016/j.jvoice.2006.03.002>
- Hill, K. T., & Miller, L. M. (2010). Auditory attentional control and selection during cocktail party listening. *Cerebral Cortex*, *20*(3), 583–590. <https://doi.org/10.1093/cercor/bhp124>
- Hirschberg, J. (2003). Pragmatics and intonation. In L. R. Horn & G. Ward (Eds.), *Handbook of pragmatics* (pp. 515–538). London, England: Blackwell.
- Hirschberg, J., & Pierrehumbert, J. (1986). The intonational structuring of discourse. *Proceedings of 24th Meeting of the Association for Computational Linguistics*, *13*, 136–144. <https://doi.org/10.3115/981131.981152>
- Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2008). Integrating cues of social interest and pitch in men's preferences for women's voices. *Biology Letters*, *4*, 192–194. <https://doi.org/10.1098/rsbl.2007.0626>
- Ladd, R. (1996). *Intonational phonology*. New York, NY: Cambridge University Press.
- Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, *50*(1), 46–70. <https://doi.org/10.1111/j.1460-2466.2000.tb02833.x>
- Lang, A. (2006). Using the limited capacity model of motivated mediated message processing to design effective cancer communication messages. *Journal of Communication*, *56*(s1), S57–S80. <https://doi.org/10.1111/j.1460-2466.2006.00283.x>
- Lang, A., Kurita, S., Rubenking, B., & Potter, R. F. (2011). miniMAM: Validating a short version of the Motivation Activation Measure. *Communication Methods & Measures*, *5*, 146–162. [https://doi.org/10.1207/s1532785xmep0102\\_4](https://doi.org/10.1207/s1532785xmep0102_4)

- Lee, H. J. (2015). *The effect of auditory pitch range on sustained and selective attention: A comparison of children with autism spectrum disorder and typically-developing children*. Open access dissertations, Paper 1396.
- Levi, S. V., & Pisoni, D. B. (2007). Indexical and linguistic channels in speech perception: Some effects of voiceovers on advertising outcomes. In T. M. Lowrey (Ed.), *Psycholinguistic phenomena in marketing communications* (pp. 203–219). Mahwah, NJ: Lawrence Erlbaum Associates.
- Liberman, M. Y., & Pierrehumbert, J. (1984). Intonational invariance under changes in pitch range and length. In M. Aronoff & R. T. Oehrle (Eds.), *Language sound structure: Studies in phonology presented to Morris Halle* (pp. 157–233). Cambridge, MA: MIT Press.
- Marie, C., & Trainor, L. (2013). Development of simultaneous pitch encoding: infants show a high voice superiority effect. *Cerebral Cortex*, *23*, 660–669. <https://doi.org/10.1093/cercor/bhs050>
- McGregor, J., & Palethorpe, S. (2008). High rising tunes in Australian English: The communicative function of  $L^*$  and  $H^*$  pitch accent onsets. *Australian Journal of Linguistics*, *28*(2), 171–193. <https://doi.org/10.1080/07268600802308766>
- Neil, E., Worrall, L., Day, A., & Hickson, L. (2003). Voice and speech characteristics and vocal hygiene in novice and professional broadcast journalists. *Advances in Speech Language Pathology*, *5*(1), 1–14. <https://doi.org/10.1080/14417040510001669001>
- Nihalani, P., & Po Lin, T. (1998). Intonation patterns in news broadcasts. *World Englishes*, *17*(1), 15–29.
- Oxenham, A. J. (2012). Pitch perception. *The Journal of Neuroscience*, *32*(39), 13335–13338.
- Perrachione, T. K., Fedorenko, E. G., Vinke, L., Gibson, E., & Dilley, L. C. (2013). Evidence for shared cognitive processing of pitch in music and language. *PLoS ONE*, *8*(8), e73372. <https://doi.org/10.1371/journal.pone.0073372>
- Pierrehumbert, J. (1980). *The phonetics and phonology of English intonation*. Doctoral dissertation. MIT, Cambridge, MA.
- Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Cohen, J. Morgan & M. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.
- Potter, R. F. (2000). The effects of voice changes on orienting and immediate cognitive overload in radio listeners. *Media Psychology*, *2*, 147–178. [https://doi.org/10.1207/S1532785XMEP0202\\_3](https://doi.org/10.1207/S1532785XMEP0202_3)
- Potter, R. F., & Bolls, P. D. (2012). *Psychophysiological measurement and meaning: Cognitive and emotional processing of media*. New York, NY: Routledge.
- Potter, R. F., Jamison-Koenig, E., Lynch, T., & Sites, J. (2016). Effect of vocal pitch difference on automatic attention to voice changes in audio messages. *Communication Research*, In press. <https://doi.org/10.1177/0093650215623835>
- Potter, R. F., Lynch, T., & Krause, A. (2015). I've heard *that* before: Habituation of the orienting response follows repeated presentation of audio structural features in radio. *Communication Monographs*, *82*, 359–378. <https://doi.org/10.1080/03637751.2015.1019529>
- Price, J. (2008). New news old news: A sociophonetic study of spoken Australian English. News broadcast speech. *AAA: Arbeiten aus Anglistik und Amerikanistik*, *33*(2), 285–310.
- Prieto, P., Payà, M., & Vanrell, M. M. (2014). Estils prosòdics en el llenguatge publicitari [Prosody styles in advertising language]. In D. Casals & J. Cornudella (Eds.), *Publicitat en*



- català. Estudis de llenguatge publicitari* [Advertising in Catalan. Studies of advertising language] (pp. 5–22). Aachen, Germany: Shaker.
- Rodero, E. (2013). The perception of a broadcasting voice. *US-China Education Review A & B*, *3*(4), 225–230.
- Rodero, E. (2015). The principle of distinctive and contrastive coherence of prosody in radio news: An analysis of perception and recognition. *Journal of Nonverbal Behavior*, *39*, 79–92. <https://doi.org/10.1007/s10919-014-0201-5>
- Rodero, E., Larrea, O., & Vázquez, M. (2013). Male and female voices in commercials. Analysis of effectiveness, adequacy for product, attention and recall. *Sex Roles*, *68*(5), 349–362. <https://doi.org/10.1007/s11199-012-0247-y>
- Selkirk, E. (2008). The relation between syntax and phonetics is indirect: the case of contrastive focus prosody. Experimental and theoretical advances in prosody. In *Cornell Prosody Conference*. Cornell, NY: UMass Amherst.
- Stephenson, A., Reese, D., & Beadle, M. (2013). *Broadcast announcing worktext*. Burlington, MA: Taylor & Francis.
- Strangert, E. (1991). Phonetic characteristics of professional news reading. *Perilus*, *13*, 39–43.
- Tracy, D. K., Ho, D. K., O'Daly, O., Michalopoulou, P., Lloyd, L. C., Dimond, E., ... Shergill, S. S. (2011). It's not what you say but the way that you say it: An fMRI study of differential lexical and non-lexical prosodic pitch processing. *BMC Neuroscience*, *12*(1), 1. <https://doi.org/10.1186/1471-2202-12-128>
- Van Lancker, D. (1980). Cerebral lateralization of pitch cues in the linguistic signal. *Research on Language & Social Interaction*, *13*(2), 201–277. <https://doi.org/10.1080/08351818009370498>
- Wells, J. C. (2007). *English intonation*. Cambridge, England: Cambridge University Press.